

# TRACKING PERFORMANCE OF A STOCHASTIC GRADIENT ALGORITHM FOR TRANSMIT ANTENNA WEIGHT ADAPTATION WITH FEEDBACK

Brian C. Banister

Department of Electrical and Computer Engineering,  
University of California at San Diego, La Jolla, CA 92093

LSI Logic Corporation, San Diego CA 92121

James R. Zeidler

Department of Electrical and Computer Engineering,  
University of California at San Diego, La Jolla, CA 92093

Space and Naval Warfare Systems Center, San Diego

## ABSTRACT

In this paper the tracking performance of a transmit antenna weight adaptation algorithm applied in a fading channel is considered. The system uses a stochastic gradient algorithm incorporating gradient sign feedback from a receiving unit to adjust the transmit weights. The feedback is simply a bit indicating which of two normalized perturbed transmission weights delivers greater power to the receiver. A reasonable performance measure is defined and an analytic estimate of the performance in an AR1 vector fading channel is derived and compared to simulations.

## 1. INTRODUCTION

It is generally accepted that the downlink of next generation cellular systems will require greater capacity than the uplink. This is largely due to the asymmetry of data traffic patterns. For example, a mobile data terminal may download large web sites while uploading only control information such as IP addresses. The use of transmit adaptive antenna arrays at the base station is a promising area for downlink capacity improvement. This paper describes a gradient algorithm utilizing mobile to base feedback in order to achieve some of those possible gains.

In many systems Frequency Division Duplexing (FDD) is used, so that the downlink and uplink channels in a multipath environment are not generally the same. This restricts how the uplink channel measured by the base station can be applied on the downlink. In general, transmit antenna algorithms in this environment can be classified as (a) space time codes, (b) blind adaptive algorithms, or (c) algorithms incorporating feedback. Space time codes achieve diversity gain by applying codes across the multiple antennas [1][3][8], and for one receive antenna provide transmit diversity without coding gain. The gains by these codes diminish as the fading channels experienced by the Tx antennas become correlated with each other, giving no gain for fully correlated fades. These algorithms perform coherent combining in the receiver, and do not achieve the gains that can arise from the array gain of adaptively weighted transmission, where the coherent combining takes place over the air. Blind algorithms utilize the measured uplink channel to infer characteristics of the downlink channel, which are then used to adapt transmit antenna weights. This may require accurate antenna calibration, as for example in estimating the angle of arrival and angular dispersion of the received signal to generate transmit weights [9], or it may assume that the long term characteristics of the uplink and downlink channels are strongly correlated [7]. However, if the antennas experience independent fading then blind techniques will not work, as without correlation between antennas no correlation from uplink to downlink channel can be extracted.

In order to benefit from both fading diversity and beam steering, an algorithm incorporating detailed downlink channel information must be used, which in FDD systems requires mobile to base feedback. This has led to several proposals for

feedback [4][5][6]. In this paper, the system proposed in [2] is considered. The usable power delivered to the mobile is considered to be an inverse cost, to be maximized, and the system proposed uses weight vector perturbations to extract a coarse estimate of the gradient of this inverse cost, which is then used to adjust the transmit antenna weights. The algorithm properties and convergence in static conditions were considered in [2]. This paper will provide an analytic performance measure for a fading channel, allowing for optimization of the adaptation rate.

## 2. PRELIMINARIES

### 2.1. System Model

The system is considered to use Nyquist pulse shaping so that a discrete time representation of the waveforms is adequate. The multipath environment is taken to introduce Rayleigh fading with only one resolvable path, which reduces the gain matrix to a rank one case. The signal received by the mobile is given by

$$r(n) = \sqrt{P^{(T)}} \frac{\mathbf{w}^H}{\|\mathbf{w}\|} \mathbf{c} \cdot s(n) + n(n) \quad (1)$$

where  $s(n)$  is the modulation sequence,  $n(n)$  is the received noise and interference,  $P^{(T)}$  is the transmit power,  $\mathbf{w}$  is the  $N \times 1$  transmit weight vector for  $N$  antennas and  $\mathbf{c}$  is the channel response vector.  $n$  is the modulation rate sample index.

With one delay path the mobile's gain matrix is of rank one.

$$\mathbf{R} \equiv \mathbf{c}\mathbf{c}^H \quad (2)$$

The total usable signal power at the receiver is given by

$$P^{(R)} = P^{(T)} \cdot \frac{\mathbf{w}^H \mathbf{R} \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \quad (3)$$

The fading channel is defined as a first order autoregressive (AR1) complex gaussian process with a zero mean complex gaussian stimulus  $\mathbf{x}$  with a time index  $i$ , slower than  $n$ .

$$\mathbf{c}(i) = a\mathbf{c}(i-1) + \mathbf{x}(i) \quad (4)$$

The channel is taken to be uncorrelated across the antennas.

$$E(\mathbf{x}(i)\mathbf{x}^H(i)) \equiv \Phi_{\mathbf{x}\mathbf{x}} \equiv 2\sigma_x^2 \mathbf{I} \quad (5)$$

### 2.2. Inverse Cost Function and Eigenanalysis

The algorithm is shown in [2] to be a stochastic gradient algorithm operating on the inverse cost function

$$J = \frac{P^{(R)}}{P^{(T)}} = \frac{\mathbf{w}^H \mathbf{R} \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \quad (6)$$

It is convenient to consider the eigenmodes of the system. Since the gain matrix (2) is rank one, the weight vector can be decomposed into two constituent eigenvectors of  $\mathbf{R}$  if properly selected. The first eigenvector is the normalized channel vector, projecting the weight vector into the range of  $\mathbf{c}$ . Since all other eigenvalues are zero the eigenvectors are arbitrary and we select the normalized projection of  $\mathbf{w}$  into the nullspace of  $\mathbf{c}$  as the

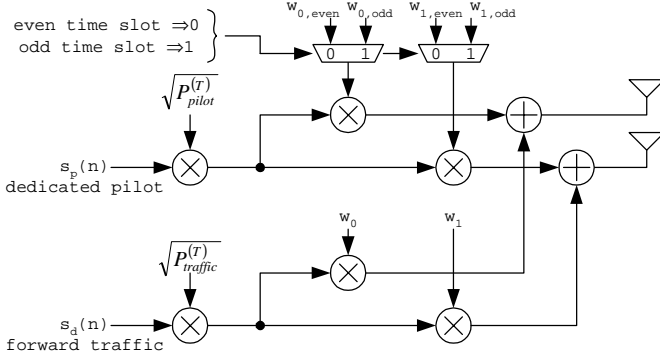


Figure 1: Example of two antenna transmitter

second eigenvector, to provide a two mode representation of  $\mathbf{w}$ .

$$\mathbf{q}_0 = \frac{\mathbf{c}}{\|\mathbf{c}\|} \quad (7)$$

$$\mathbf{q}_1 = \left( \mathbf{I} - \frac{\mathbf{c}\mathbf{c}^H}{\|\mathbf{c}\|^2} \right) \mathbf{w} / \left\| \left( \mathbf{I} - \frac{\mathbf{c}\mathbf{c}^H}{\|\mathbf{c}\|^2} \right) \mathbf{w} \right\| \quad (8)$$

Then the weight vector is portrayed in terms of these two eigenvectors as

$$\mathbf{u} = \mathbf{Q}^H \mathbf{w} \quad (9)$$

$\begin{matrix} 2 \times 1 & 2 \times N & N \times 1 \end{matrix}$

It will be convenient to define the vector  $\mathbf{v}$  comprised of the squared magnitudes of  $\mathbf{u}$ . I.e.,  $v_0$  is the “energy” of  $\mathbf{w}$  in the channel subspace, which is receivable by the mobile, and  $v_1$  is the error “energy” of  $\mathbf{w}$  in the channel nullspace, which cannot be received by the mobile.

$$[\mathbf{v}]_n \equiv |u_n|^2 \Rightarrow \mathbf{v} = \begin{bmatrix} \frac{|\mathbf{c}^H \mathbf{w}|^2}{\|\mathbf{c}\|^2} & \left\| \left( \mathbf{I} - \frac{\mathbf{c}\mathbf{c}^H}{\|\mathbf{c}\|^2} \right) \mathbf{w} \right\|^2 \end{bmatrix}^T \quad (10)$$

Note that

$$J = \frac{v_0}{v_0 + v_1} \cdot \|\mathbf{c}\|^2 \quad (11)$$

### 3. ADAPTATION BY GRADIENT SIGN FEEDBACK

#### 3.1. Algorithm Description

The system has been described in more detail in [2]. The nomenclature will reflect pilot assisted coherent CDMA. The transmitter transmits the data with a weight vector  $\mathbf{w}$ , while the weight vector applied to the pilot alternates between odd and even perturbed values as shown in Figure 1.

$$\mathbf{t}(n) = \sqrt{P_{\text{traffic}}^{(r)}} \frac{\mathbf{w}}{\|\mathbf{w}\|} s_d(n) + \sqrt{P_{\text{pilot}}^{(r)}} s_p(n) \cdot \begin{cases} \frac{\mathbf{w}_{\text{even}}}{\|\mathbf{w}_{\text{even}}\|} & \text{even timeslots} \\ \frac{\mathbf{w}_{\text{odd}}}{\|\mathbf{w}_{\text{odd}}\|} & \text{odd timeslots} \end{cases} \quad (12)$$

A measurement and feedback interval can be comprised of several even/odd time slots. The weights are generated from a tracked “base” weight vector  $\mathbf{w}_{\text{base}}$  algorithm as follows.

$$\mathbf{w}_{\text{even}} = \mathbf{w}_{\text{base}} + \|\mathbf{w}_{\text{base}}\| \beta \sqrt{\frac{\pi}{2}} \cdot \mathbf{p} \quad (13)$$

$$\mathbf{w}_{\text{odd}} = \mathbf{w}_{\text{base}} - \|\mathbf{w}_{\text{base}}\| \beta \sqrt{\frac{\pi}{2}} \cdot \mathbf{p} \quad (14)$$

$$\mathbf{w} = \mathbf{w}_{\text{even}} + \mathbf{w}_{\text{odd}} \quad (15)$$

$\mathbf{p}$  is a zero mean complex gaussian perturbation vector and

$$E(\mathbf{p}\mathbf{p}^H) = 2\mathbf{I} \quad (16)$$

The receiver provides a feedback bit selecting which time slot, even or odd, provided a larger delivered power and the transmitter adjusts the weights by updating  $\mathbf{w}_{\text{base}}$  to the value of the associated perturbed vector. The perturbation vector and transmission weight vectors are regenerated and the process continues. As shown in [2], this is a stochastic gradient adaptation algorithm operating on the inverse cost  $J$  (eq. (6)), and the algorithm parameter  $\beta$  is the adaptation rate. Setting the data channel weight vector to the sum of the even and odd pilot vectors allows for simple coherent reception, as an efficient receiver can use the mean of the even and odd time slot channel estimates as the channel estimate for demodulation of the data.

#### 3.2. Dynamic Performance

The performance of the algorithm will be considered by deriving expressions for a transition matrix applied to the eigenmodal energy vector  $\mathbf{v}$  of the weight vector  $\mathbf{w}$ . This transition matrix incorporates the effect of both the algorithm update and the channel change for one time step. Each step is shown to be a linear transformation of the vector  $\mathbf{v}$  using the following approximations: (1) independence of the channel vector norm from update to update, (2) second order Taylor approximation of the fading channel update on the weight response in the range of  $\mathbf{c}$ , i.e.  $v_0$ . The first assumption will more closely apply for large numbers of antennas, as the channel norm approaches a constant. The second is found to be reasonable because the expectation of odd order terms are zero, and the 4<sup>th</sup> order term will be small.

The step update from the algorithm feedback decision was derived in [2]. Translating that result into a simplified notation for this rank one situation (only 1 resolvable path) with

$$\gamma(i) = \sqrt{\frac{v_0(i)}{v_1(i)}} \quad (17)$$

the algorithm update is given by

$$\mathbf{G}_{\text{alg}}(i) = (1 - 2\beta\gamma(i))\mathbf{I} + 2\beta\left(\gamma(i) + \frac{1}{\gamma(i)}\right) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \beta^2 \pi \begin{bmatrix} 1 & 1 \\ N-1 & N-1 \end{bmatrix} \quad (18)$$

$$\tilde{\mathbf{v}}(i+1) = \mathbf{G}_{\text{alg}}(i)\mathbf{v}(i) \quad (19)$$

where the vector  $\mathbf{v}$  contains the magnitude squared of the projection of the weight vector into the range of  $\mathbf{c}$  and into the null space of  $\mathbf{c}$ , and  $\gamma$  is the ratio of the magnitudes of the elements of  $\mathbf{u}$ .  $\tilde{\mathbf{v}}$  is an intermediate value, before the effect of the change of the fading channel is considered. For the transition of  $\mathbf{v}$  due to the time varying channel, a 2<sup>nd</sup> order Taylor approximation of the incremental step due to the new stimulus  $\mathbf{x}$  is applied, given by the gradient vector and Hessian matrix of  $v_0$  with respect to  $\mathbf{c}$ . The effect of the channel change is approximately (time indices for  $\mathbf{w}$ ,  $\mathbf{x}$ ,  $\mathbf{c}$  are omitted as they become irrelevant using the approximations)

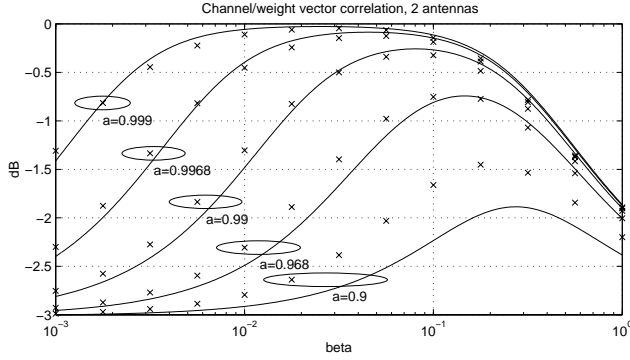


Figure 2: Mean correlation  $v_0$ ,  $N=2$ , simulated and from (26), (27)

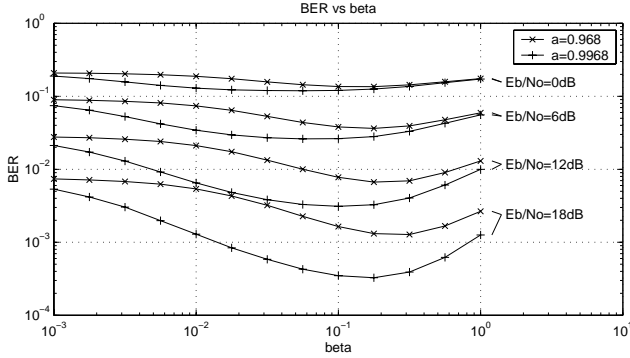


Figure 3: Bit error rates for  $N=2$ ,  $a=[0.968, 0.9968]$

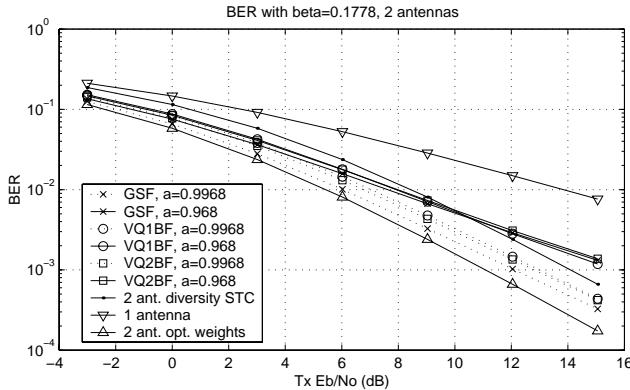


Figure 4: Simulated BER for  $N=2$ ,  $\beta=0.1778$ ,  $a=[0.968, 0.9968]$ , compared to optimal weights, STC, vector quantization

$$\begin{aligned}
 v_0(i+1) \equiv & \tilde{v}_0(i+1) + \frac{2}{a\|\mathbf{c}\|^2} \text{Re}(\mathbf{x}^H \mathbf{w} \mathbf{w}^H \mathbf{c} - \tilde{v}_0(i+1) \mathbf{x}^H \mathbf{c}) \\
 & + \frac{1}{a^2\|\mathbf{c}\|^2} \left( |\mathbf{x}^H \mathbf{w}|^2 - 4 \text{Re} \left( \frac{\mathbf{c}^H \mathbf{w}}{\|\mathbf{c}\|^2} \mathbf{x}^H \mathbf{c} \cdot \mathbf{w}^H \mathbf{x} \right) \right. \\
 & \left. + 4 \frac{\tilde{v}_0(i+1)}{\|\mathbf{c}\|^2} |\mathbf{x}^H \mathbf{c}|^2 - \tilde{v}_0(i+1) \|\mathbf{x}\|^2 \right)
 \end{aligned} \quad (20)$$

Taking the expectation of (20) with respect to both  $\mathbf{x}$  and  $\mathbf{c}$  (using assumption 1) gives, with the ones vector  $\mathbf{1}$

$$s \equiv E \left( \frac{2\sigma_x^2}{a^2\|\mathbf{c}\|^2} \right) = \frac{(1-a^2)}{a^2 \cdot (N-1)} \quad (21)$$

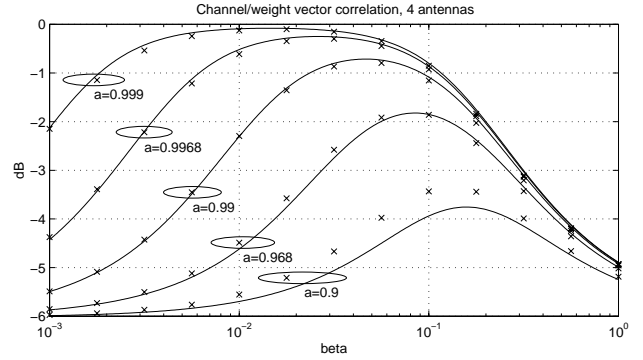


Figure 5: Mean correlation  $v_0$ ,  $N=4$ , simulated and from (26), (27)

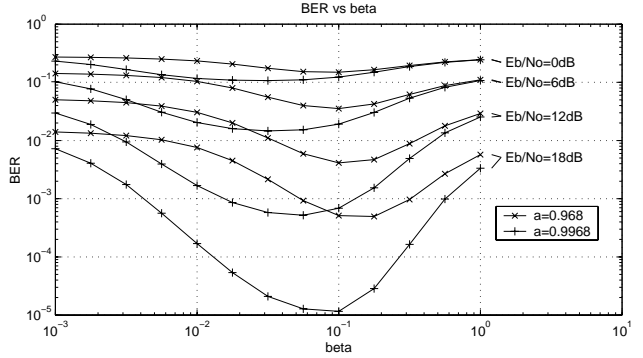


Figure 6: Bit error rates for  $N=4$ ,  $a=[0.968, 0.9968]$

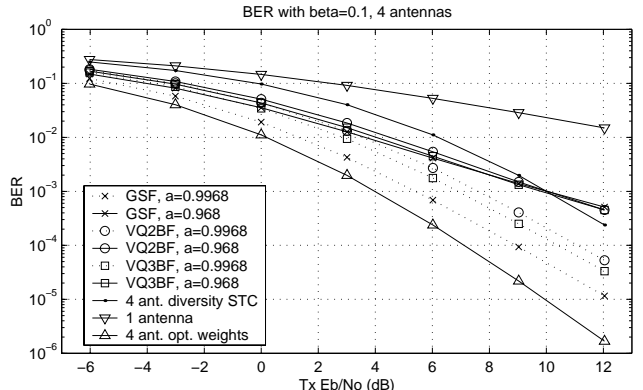


Figure 7: Simulated BER for  $N=4$ ,  $\beta=0.1$  and  $a=[0.968, 0.9968]$ , compared to optimal weights, STC, vector quantization

$$v_0(i+1) \equiv \tilde{v}_0(i+1) + s \cdot (\mathbf{1}^T \tilde{\mathbf{v}}(i+1) - N \tilde{v}_0(i+1)) \quad (22)$$

Equation (22) will be used as an approximation to the update in the desired eigenspace. Given this update in  $v_0$ , the update in  $v_1$  is known, as the weight norm is not changed by a change in the channel. Applying the conservation of the weight vector norm gives the channel update transition.

$$\mathbf{v}(i+1) \equiv \mathbf{G}_{\text{chan}} \tilde{\mathbf{v}}(i+1) = \left( \mathbf{I} + s \cdot \begin{bmatrix} 1-N & 1 \\ N-1 & -1 \end{bmatrix} \right) \tilde{\mathbf{v}}(i+1) \quad (23)$$

Note that the update of (23) is intuitively satisfying. For small  $\sigma_x$  the update represents a transfer of an equal fraction of modal power from each of the  $N$  eigenmodes to all  $N$  eigenmodes. The  $N-1$  weighting arises from the representation of all  $N-1$  nullspace eigenmodes in the single term  $v_1$ . Clearly the approximation of (22) can only be valid if the resultant values

of  $v_0$  remain positive, so we find that the AR1 rate of change must be slow enough to satisfy

$$1 - a \ll 1 - \frac{1}{\sqrt{2}} \quad (24)$$

The step update for both the algorithm and channel change is then approximated as

$$\mathbf{v}(i+1) = \mathbf{G}_{\text{chan}} \mathbf{G}_{\text{alg}}(i) \mathbf{v}(i) \equiv \mathbf{G}(i) \mathbf{v}(i) \quad (25)$$

The transition matrix  $\mathbf{G}(i)$  incorporates first the algorithm update based on the measured channel at time  $i$  and then the channel update to time  $i+1$ . It can be shown that with iterative application of  $\mathbf{G}(i)$  as in (25),  $\gamma$  converges to a steady state value

$$\gamma|_{\text{steady state}} = \frac{2\beta - 4N\beta s + \sqrt{4(N-1)\beta^4\pi^2 + 4\beta^2 + 16N^2\beta^2s^2 - 16N\beta^2s + 16(N-1)\beta^2\pi s + 16(N-1)s^2}}{2(\beta^2\pi + 2s)(N-1)} \quad (26)$$

This gives the steady state solution for the “expected” normalized correlation value  $v_0$  by using

$$v_0|_{\text{steady state}} = \lim_{i \rightarrow \infty} \left( \frac{v_0(i)}{v_0(i) + v_1(i)} \right) = \frac{\gamma^2|_{\text{steady state}}}{1 + \gamma^2|_{\text{steady state}}} \quad (27)$$

### 3.3. Discussion and numerical results

Equation (26) provides a mechanism for evaluating the effectiveness of an adaptation rate  $\beta$  for a fading rate of the AR1 process. It is of interest to note that with  $\beta=0$

$$\gamma^2|_{\text{steady state}, \beta=0} = \frac{1}{(N-1)} \quad (28)$$

This is a confirmation of the common sense solution in this case, where  $1/N$  of the weight vector energy lies within span of the channel vector and the remainder is in its nullspace.

To confirm the applicability of this analysis simulations were performed for  $N=2$  and  $N=4$  antennas with fading uncorrelated across the antennas as per equation (5), over a variety of values of the AR1 parameter  $a$  and the adaptation rate  $\beta$ . In addition, bit error rates were simulated in order to compare the BER performance with the performance of the vector “cross correlation”  $v_0$ . The simulations implemented the receiver’s decision and the feedback channel with no errors.

The simulation results are compared to the analysis in Figure 2 and Figure 5 for 2 and 4 antennas respectively. Both figures show that the simulated vector correlation has a very good match with the analysis for all fading rates except  $a=0.9$ , for which we note that the AR process is approaching the limit of (24). Comparing these figures to Figure 3 and Figure 6 we see that the analysis provides a good prediction of the best value of  $\beta$  for minimizing bit error rates.

Bit error rates from the simulation of the gradient sign feedback (GSF) are shown for two of the fading rates in Figure 4 and Figure 7. For comparison, these figures also include the analytic performance of diversity space time codes (STC) for a single Rx antenna (with no coding gain, as in [1] for two Tx antennas) and simulated performance of vector quantization code book selection feedback, wherein the receiver provides feedback selecting which of several weight vectors is best [5]. For two antennas, VQ1BF is one bit feedback selecting which antenna should transmit ( $2^{\text{nd}}$  order selection diversity), VQ2BF is two bit feedback selecting a phase rotation of  $\langle 0, \pi/2, \pi, 3\pi/2 \rangle$  for the second weight. For 4 antennas, VQ2BF is two bit

feedback selecting which of the 4 antennas should transmit ( $4^{\text{th}}$  order selection diversity) and VQ3BF is 3 bit feedback selecting a phase rotation of  $\langle 0, \pi \rangle$  for weights  $w_1, w_2, w_3$ . For two or three bit feedback, the feedback decision interval is lengthened so that the feedback data rate is unchanged.

BERs are plotted versus Tx  $E_b/N_0$  normalized to the Rx  $E_b/N_0$  for 1 Tx antenna, so that the 3.01dB and 6.02dB array gain from 2 antenna or 4 antenna systems can be seen. For slow fading the algorithm performs close to the theoretic limit and outperforms the vector quantization feedback approaches. For faster fading, the gradient approach gives similar performance to vector quantization, but both feedback approaches are outperformed by diversity space time coding, which does not require the transmitter to adapt to the time varying channel.

## 4. CONCLUSION

An analysis of the tracking performance of the gradient sign feedback implementation of transmit antenna array weight adaptation has been presented. The performance is considered for the case of a first order autoregressive Rayleigh fading process. The analysis is verified through simulation. For slow fading channels the algorithm is found to give a better BER than diversity space time coding and vector quantization feedback.

## 5. REFERENCES

- [1] S. Alamouti “A Simple Transmit Diversity Technique for Wireless Communications”, *IEEE Journal on Select Areas in Communications*, Vol. 16, No.8, pp 1451-1458, Oct. 1998.
- [2] B.C. Banister, J.R. Zeidler “A Stochastic Gradient Algorithm for Transmit Antenna Weight Adaptation with Feedback”, to be published *Proc. 3<sup>rd</sup> IEEE Workshop on Signal Processing Advances in Wireless Communications*, March 2001
- [3] G. J. Foschini “Layered Space Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas”, *Bell Labs Technical Journal*, pp 41-59, Autumn 1996
- [4] D. Gerlach, A. Paulraj “Adaptive Transmitting Antenna Arrays with Feedback”, *IEEE Signal Processing Letters*, Vol. 1, No. 10, pp 150-152, Oct. 1994.
- [5] R.W. Heath Jr., A. Paulraj, “A Simple Scheme for Transmit Diversity Using Partial Channel Feedback”, *Conference Record of Thirty Second Asilomar Conference on Signals, Systems and Computers*, pp 1073-1078, Nov. 1998.
- [6] J.W. Liang, A. Paulraj “Forward link antenna diversity using feedback for indoor communication systems” *Proceedings, 1995 International Conference on Acoustics, Speech and Signal Processing*, pp 1753-1755, May 1995.
- [7] G.G. Raleigh, S.N. Diggavi, V.K. Jones, A. Paulraj “A Blind Adaptive Transmit Antenna Algorithm for Wireless Communication”, *Proceedings IEEE International Conference on Communications*, pp 1494-1499, June 1995
- [8] V. Tarokh, H. Jafarkhani, A.R. Calderbank, “Space-Time Block Coding for Wireless Communications: Performance Results” *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 3, pp 451-460, March 1999
- [9] P. Zetterberg, B. Ottersten, “The Spectrum Efficiency of a Base Station Antenna Array System for Spatially Selective Transmission”, *IEEE Transactions on Vehicular Technology*, Vol. 44, No. 3, pp 651-660, August 1995.